JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 53, 359–366 (1976)

On Moment-Discretization and Least-Squares Solutions of Linear Integral Equations of the First Kind

M. Z. NASHED

School of Mathematics, Georgia Institute of Technology, Atlanta, Georgia 30332

Submitted by Kenneth L. Cooke

Let $K(s, t)$ be a continuous function on $[0, 1] \times [0, 1]$, and let $\mathscr{K}$ be the linear integral operator induced by the kernel $K(s, t)$ on the space $\mathscr{L}_2[0, 1]$. This note is concerned with moment-discretization of the problem of minimizing $\| \mathscr{K}x - y \|$ in the $\mathscr{L}_2$-norm, where y is a given continuous function. This is contrasted with the problem of least-squares solutions of the moment-discretized equation: $\int_0^1 K(s_i, t)\, x(t)\, dt = y(s_i)$, $i = 1, 2,..., n$. A simple commutativity result between the operations of "moment-discretization" and "least-squares" is established. This suggests a procedure for approximating $\mathscr{K}^\dagger y$ (where $\mathscr{K}^\dagger$ is the generalized inverse of $\mathscr{K}$), without recourse to the normal equation $\mathscr{K}^*\mathscr{K}x = \mathscr{K}^*y$, that may be used in conjunction with simple numerical quadrature formulas plus collocation, or related numerical and regularization methods for least-squares solutions of linear integral equations of the first kind.

1. INTRODUCTION AND PRELIMINARIES

We consider the integral equation of the first kind,

$$(\mathscr{K}x)(s) := \int_0^1 K(s, t)\, x(t)\, dt = y(s), \qquad 0 \leqslant s \leqslant 1, \tag{1}$$

where the kernel $K(x, t)$ is continuous on $[0, 1] \times [0, 1]$ and y is a given continuous function on $[0, 1]$. Let $(u_n, v_n; \mu_n)$ be a singular system of $K(s, t)$, i.e.,

$$u_n = \mu_n \mathscr{K} v_n, \qquad v_n = \mu_n \mathscr{K}^* u_n, \tag{2}$$

where $\mathscr{K}^*$ is the adjoint of $\mathscr{K}$ and $\{u_n\}$, $\{v_n\}$ are orthonormal systems of functions in $\mathscr{L}_2[0, 1]$. It is well known [16, pp. 164–166] that Eq. (1) has a solution $x(t)$ in $\mathscr{L}_2[0, 1]$ if and only if

$$\langle y, u \rangle = 0 \qquad \text{for every } u \in \mathscr{L}_2[0, 1] \quad \text{such that} \quad K^*u = 0, \tag{3}$$

and

$$\sum_{n=1}^{\infty} \mu_n^2 \, |\langle y, u_n \rangle|^2 < \infty. \tag{4}$$

(Here, $\langle \, , \rangle$ denotes the inner product in $\mathscr{L}_2[0, 1]$). Conditions (3) and (4) are also known as Picard's criterion. It follows from (2) that u_n and μ_n^2 are the orthonormalized eigenfunctions and the corresponding eigenvalues, respectively, of the compact self-adjoint linear operator $\mathscr{K}\mathscr{K}^*$. Condition (4) forces an element $y \in \mathrm{cl}\,\mathscr{R}(\mathscr{K})$, the closure of the range of $\mathscr{K}$, to belong to the range of $\mathscr{K}$. For equivalent manifestations of this criterion, see [4]. Since $\mathscr{R}(\mathscr{K})$ for a compact linear operator $\mathscr{K}$ is not closed (unless $\mathscr{R}(\mathscr{K})$ is a finite dimensional subspace), Picard's criterion for existence of a solution of an integral equation of the first kind is more complicated than its counterpart for an integral equation of the second kind (where the range of $I-\lambda\mathscr{K}$ is always closed); the usual Fredholm alternative theorem does not hold in the former case.

When $y \notin \mathscr{R}(\mathscr{K})$, one may resort to the useful notion of a *least-squares solution* u of (1) defined by

$$\| \mathscr{K}u - y \| = \inf\{\| \mathscr{K}x - y \| : x \in \mathscr{L}_2[0, 1]\} \tag{5}$$

(the norm here is that induced by the inner product on $\mathscr{L}_2[0, 1]$). It is easy to show, since $\mathscr{R}(\mathscr{K})$ is not closed, that (1) does not have a least-squares solution if the orthogonal projection of y on the closed subspace $\mathrm{cl}\,\mathscr{R}(\mathscr{K})$ is not in $\mathscr{R}(\mathscr{K})$. Moreover, Eq. (1) has a least-squares solution for all $y \in \mathscr{R}(\mathscr{K}) \oplus \mathscr{R}(\mathscr{K})^\perp$. The *generalized inverse* of $\mathscr{K}$, denoted by $\mathscr{K}^\dagger$, is the (linear) operator whose maximal domain is $\mathscr{D}(\mathscr{K}^\dagger) = \mathscr{R}(\mathscr{K}) \oplus \mathscr{R}(\mathscr{K})^\perp$ and which associates with each $y \in \mathscr{D}(\mathscr{K}^\dagger)$ the unique least-squares solution of (1) of minimal norm. We call $\mathscr{K}^\dagger y$, for $y \in \mathscr{D}(\mathscr{K}^\dagger)$, the pseudosolution of (1). For each $y \in \mathscr{D}(\mathscr{K}^\dagger)$, the set S_y of all least-squares solutions of (1) is given by

$$S_y = \mathscr{K}^\dagger y \oplus \mathscr{N}(\mathscr{K}), \tag{6}$$

where $\mathscr{N}(\mathscr{K})$ is the null space of $\mathscr{K}$. For an exposition on generalized inverse of linear operators, see [10].

Iterative methods (such as successive approximations, steepest descent, and conjugate gradient methods) for least-squares solutions of (1) in $\mathscr{L}_2[0, 1]$ converge as $(1/n)$ (see [8, 10]). Furthermore, since $\mathscr{K}^\dagger$ in this case is unbounded and densely defined, the problem (1) is illposed; the pseudosolution $\mathscr{K}^\dagger y$ does not depend continuously on y for $y \in \mathscr{R}(\mathscr{K}) \oplus \mathscr{R}(\mathscr{K})^\perp$ and does not exist at all for $y \notin \mathscr{R}(\mathscr{K}) \oplus \mathscr{R}(\mathscr{K})^\perp$. Hence, one has to resort to some "regularization" schemes and/or confine attention to specific approximation methods, that overcome, under restrictions on the class of admissible solutions, the inherent ill-posedness and numerical instability of this problem. See [7, 11, 12, 17, 18] for some of these methods and for further references. The inevitable moment discretization or numerical quadrature plus collocation occurs at some stage in many of these methods.

This note is concerned with the question of "commutativity" of moment-discretization and least-squares in this setting.

2. Least-Squares Solutions of Moment Discretization of $\mathscr{K}x = y$

Suppose now that $y(s)$ is known at a finite number of points

$$0 \leqslant s_1 < s_2 < \cdots < s_n \leqslant 1,$$

and consider the moment-discretization of (1), namely,

$$\int_0^1 K(s_i, t)\, x(t)\, dt = y(s_i), \qquad i = 1,..., n. \tag{7}$$

Let $K_i := K(s_i, \ )$, $y_i := y(s_i)$, and rewrite (7) in the form $\langle x, \bar{K}_i \rangle = y_i$, $i = 1,..., n$, where the bar denotes conjugate. Define the operator $T_n: \mathscr{L}_2[0, 1] \to \mathbb{R}^n$ by

$$T_n x = \begin{pmatrix} \langle x, \bar{K}_1 \rangle \\ \langle x, \bar{K}_2 \rangle \\ \vdots \\ \langle x, \bar{K}_n \rangle \end{pmatrix} \qquad \text{and} \qquad Y_n = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}.$$

Then (7) takes the form

$$T_n x = Y_n . \tag{8}$$

Since $\mathscr{R}(T_n)$ is a finite-dimensional space, (8) always has least-squares solutions, the generalized inverse $T_n^{\dagger}$ exists on $\mathbb{R}^n$, and the set of all least-squares solutions of (8) is given by

$$\mathscr{S}_y^{(n)} = T_n^{\dagger} Y_n \oplus \mathscr{N}(T_n). \tag{9}$$

The set $\mathscr{S}_y^{(n)}$, defined by (9), coincides with the set of all solutions u of the equation

$$T_n^* T_n u = T_n^* Y_n , \tag{10}$$

where T_n^*, the adjoint of T_n, is given by

$$T_n^* v = \sum_{i=1}^{n} v_i \bar{K}(s_i, \ ), \tag{11}$$

where $v = (v_1 ,..., v_n)'$ and the prime denotes transpose.

To verify (11), note that

$$\langle T_n u, v\rangle = \sum_{i=1}^{n} \langle u, \bar{K}_i\rangle \bar{v}_i = \sum_{i=1}^{n} \bar{v}_i \int_0^1 K(s_i, t)\, u(t)\, dt,$$

and

$$\langle u, T_n^* v\rangle = \int_0^1 u(t)\, \overline{T_n^* v}\, dt = \int_0^1 \left[u(t) \sum_{i=1}^{n} \bar{v}_i K(s_i, t)\right] dt = \langle T_n u, v\rangle.$$

Now $T_n^* T_n: \mathscr{L}_2[0, 1] \to \mathscr{L}_2[0, 1]$. However, $\mathscr{R}(T_n^* T_n)$ is a finite-dimensional subspace of $\mathscr{L}_2[0, 1]$; in fact,

$$\mathscr{R}(T_n^* T_n) = \mathscr{R}(T_n^*) = \operatorname{span}\{\bar{K}_1, ..., \bar{K}_n\}. \tag{12}$$

Equation (8) has a unique least-squares solution x_n^* in $\mathscr{R}(T_n^*)$, namely, the least-squares solution of minimal $\mathscr{L}_2$-norm:

$$x_n^*(t) = T_n^\dagger Y_n(t) = \sum_{i=1}^{n} \alpha_i \bar{K}(s_i, t). \tag{13}$$

As we shall see below, x_n^* is a solution of a *consistent* linear integral equation with a degenerate kernel.

In particular, if $y \in \mathscr{R}(\mathscr{K})$, then substituting the right-hand side of (13) in (8), one easily finds that x_n^* is given explicitly by

$$x_n^*(\ ) = (y_1, y_2, ..., y_n)\, Q_n^\dagger (\bar{K}_1, \bar{K}_2, ..., \bar{K}_n)', \tag{14}$$

where Q_n is the $n \times n$ matrix whose *ij*th element is given by

$$Q(s_i, s_j) = \int_0^1 \bar{K}(s_i, t)\, K(s_j, t)\, dt.$$

Here $Q_n^\dagger$ is the Moore–Penrose generalized inverse of Q_n (see [1, 10, 14]).

All least-squares solutions of (8) are of the form $x_n = T_n^\dagger Y_n + Pz$, where z is any element in $\mathscr{L}_2[0, 1]$ and P is the orthogonal projector of $\mathscr{L}_2[0, 1]$ onto $\mathscr{N}(T)$; in other words, $x_n = x_n^* + w$, where $w \in [\operatorname{span}\{\bar{K}_i\}_{i=1}^n]^\perp$.

Substituting (11) in (10), we obtain

$$\sum_{i=1}^{n} \langle u, \bar{K}_i\rangle \bar{K}(s_i, t) = \sum_{i=1}^{n} y_i \bar{K}(s_i, t),$$

or

$$\int_0^1 \left\{\sum_{i=1}^{n} K(s_i, r)\, \bar{K}(s_i, t)\right\} u(r)\, dr = \sum_{i=1}^{n} y_i \bar{K}(s_i, t). \tag{15}$$

Note that this is an integral equation of a first kind with a degenerate kernel. Moreover Eq. (15) is always consistent, even though Eq. (1), which we started from, may be inconsistent.

Moment discretization of (15) on the grid $\{t_1, ..., t_n\}$ yields

$$\sum_{i=1}^{n} \left\{\int_0^1 K(s_i, r)\, u(r)\, dr\right\} \bar{K}(s_i, t_j) = \sum_{i=1}^{n} y(s_i)\, \bar{K}(s_i, t_j), \qquad j = 1, 2, ..., n. \tag{16}$$

Note in particular that if the set $\{K_i\}_{i=1}^n$ is linearly independent, then the least-squares solutions of (8) are, in fact, solutions and $T_n{}^\dagger Y_n$ is the unique solution of minimal-norm. Furthermore, the α_i's in (13) are uniquely determined from the system $\sum_{j=1}^n \langle K_j, K_i\rangle \alpha_j = y_i$, $i = 1, ..., n$, since the Gram matrix $[\langle K_i, K_j\rangle]$ is nonsingular in this case, so that $Q_n{}^\dagger$ may be replaced by Q_n^{-1} in (14).

3. Moment Discretization of the Normal Equation $\mathscr{K}^*\mathscr{K}x = \mathscr{K}^*y$

Equation (7) represents the moment discretization of (1) and is quite appropriate if $y \in \mathscr{R}(\mathscr{K})$. However, if $y \notin \mathscr{R}(\mathscr{K})$ but $y \in \mathscr{D}(\mathscr{K}^\dagger)$, then (1) has a least-squares solution that is, equivalently, a solution of the equation

$$\mathscr{K}^*\mathscr{K}x = \mathscr{K}^*y. \tag{17}$$

Moment discretization of the least-squares problem (5) when $y \notin \mathscr{R}(\mathscr{K})$ should, strictly speaking, be applied to Eq. (17) rather than (1). Noting that

$$\mathscr{K}^*\mathscr{K}x = \int_0^1 M(\ , t)\, x(t)\, dt,$$

where

$$M(s, t) = \int_0^1 \bar{K}(r, s)\, K(r, t)\, dr,$$

we may write (17) in the form

$$\int_0^1 \int_0^1 \{\bar{K}(r, s)\, K(r, t)\, dr\}\, x(t)\, dt = \int_0^1 \bar{K}(t, s)\, y(t)\, dt.$$

For $0 \leqslant s_1 < s_2 < \cdots < s_n \leqslant 1$, we have $\langle M_i, \bar{x}\rangle = (\mathscr{K}^*y)(s_i)$, or

$$\int_0^1 \int_0^1 \bar{K}(r, s_i)\, K(r, t)\, x(t)\, dr\, dt = \int_0^1 \bar{K}(t, s_i)\, y(t)\, dt, \qquad i = 1, ..., n. \tag{18}$$

Thus, if x is a least-squares solution of (1), then x is a solution of (18). We rewrite (18) as

$$\int_0^1 \left\{\int_0^1 K(r, t)\, x(t)\, dt\right\} \bar{K}(r, s_i)\, dr = \int_0^1 \bar{K}(r, s_i)\, y(r)\, dr, \qquad i = 1,\ldots, n. \tag{19}$$

Now we discretize the integrals with respect to r and at this stage use, for simplicity, *uniform mesh* size and also the *same mesh* for *all* the preceding discretizations. Then (18) becomes

$$\sum_{j=1}^{n} \left\{\int_0^1 K(r_j, t)\, x(t)\, dt\right\} \bar{K}(r_j, s_i) = \sum_{j=1}^{n} \bar{K}(r_j, s_i)\, y(r_j) \qquad i = 1,\ldots, n,$$

which is the same system as (16).
Thus, we have the commutative diagram

$$\begin{tikzcd}
\mathcal{K}x = y \arrow[r, "\text{Least squares}"] \arrow[d, "s=s_i"] & \mathcal{K}^*\mathcal{K}x = \mathcal{K}^*y \arrow[r, "\text{Moment}", "\text{Discretization}"'] & (\mathcal{K}^*\mathcal{K}x)\, s_i = (\mathcal{K}^*y)\, s_i \arrow[dd, "\text{Approximate integral by sum over uniform mesh}"'] \\
\langle \mathcal{K}_i, x\rangle = y_i \arrow[d, "\text{Least squares}"] & & \\
T_n{}^*T_n x_n = T_n{}^*Y_n \arrow[rr, "\text{Moment}", "\text{Discretization}"'] & & \tilde{x}_n(t) = \hat{x}_n(t)
\end{tikzcd}$$

Remark 1. The assumption of uniform and identical mesh sizes led to the commutativity relation $\tilde{x}_n(t) = \hat{x}_n(t)$ (see the diagram) for ***each n***. In practice it is more desirable to choose the s_i most densely where the function y is changing most rapidly, based on the experimentally observed values of $y(s)$. In this case the commutativity relation does not necessarily hold for each n; however, it remains valid asymptotically. Full discretization may be used to reduce the system of linear relations (16) to a system of matrix equations without disturbing the commutativity relation. In particular, one may use minimum-variance formulas that take into consideration the inherent error in the observed values of the functions (see, for example, [5]).

Remark 2. The implication of the commutativity relation for numerical solutions of the least-squares problem (5) is immediate. One may follow the route of Section 2, applying computational methods for generalized inverses of matrices, and arrive at an approximation to $\hat{x}(t) = (\mathscr{K}^{\dagger}y)(t)$. (For a comparison of some direct methods for computing generalized inverses, see [15, 18] and its corresponding bibliography and the exposition and references in [1, 2, 14]. For the use of generalized inverses in matrix equations arising from discretization of differential and integral operators, see also [9].)

Remark 3. Generalization to operator equations and abstract discretization schemes [3, 6, 13] is possible in a similar way as, for instance, in [13] and may be used in connection with a variety of moment discretization problems [6]. In [13] a theory of discretization processes is developed that proves that for a wide class of operators, discretization and differentiation indeed "commute." As a by-product of this, Ortega and Rheinboldt [13] formalized the observation that in the numerical solution of operator equations, discretization followed by Newton's method results in the same linear algebraic system of equations as application of Newton's method to the operator equation followed by discretization. Rates of convergence of moment-discretization of least-squares solutions of integral and operator equations have been obtained recently by Nashed and Wahba [11].

Acknowledgment

This work was sponsored by the United States Army under Contract No. DA-31-124-ARO-D-462.

References

1. A. Albert, "Regressions and the Moore–Penrose Pseudoinverse," Academic Press, New York, 1972.
2. A. Ben-Israel and T. N. E. Greville, "Generalized Inverses: Theory and Applications," Wiley (Interscience), New York, 1974.
3. J. W. Daniel, "The Approximate Minimization of Functions," Prentice–Hall, Englewood Cliffs, N.J., 1971.
4. J. B. Diaz and F. T. Metcalf, On iteration procedures for equations of the first kind, $Ax = y$, and Picard's criterion for the existence of a solution, *Math. Comp.* **24** (1970), 923–935.
5. C. S. Duris, Optimal quadrature formulas using generalized inverses. Part I. General theory and minimum variance formulas, *Math. Comp.* **25** (1971), 495–504.
6. M. Golomb and H. F. Weinberger, Optimal approximation and error bounds, *in* "On Numerical Approximation" (Proc. Sympos., Madison, Wis., 1958) (R. E. Langer, Ed.), pp. 117–190, Univ. of Wisconsin Press, Madison, 1959.
7. R. J. Hanson, A numerical method for solving Fredholm integral equations of the first kind, *SIAM J. Numer. Anal.* **8** (1971), 616–662.
8. W. J. Kammerer and M. Z. Nashed, Iterative methods for best approximate solutions of linear integral equations of the first and second kinds, *J. Math. Anal. Appl.* **40** (1972), 547–573.

9. A. Korganoff and M. Pavel-Parvu, "Méthodes de calculs numérique-2. Éléments de théorie des matrices carrées, et rectangles en analyse numérique," Dunod, Paris, 1967.
10. M. Z. Nashed, Generalized inverses, normal solvability, and iteration for singular operator equations, *in* "Nonlinear Functional Analysis and Applications" (L. B. Rall, Ed.), pp. 311–359, Academic Press, New York, 1971.
11. M. Z. Nashed and Grace Wahba, Convergence rates of approximate least squares solutions of linear integral and operator equations of the first kind, *Math. Comp.* **28** (1974), 69–80.
12. M. Z. Nashed and Grace Wahba, Approximate regularized solutions to linear operator equations when the data-vector is not in the range of the operator, to appear.
13. J. M. Ortega and W. C. Rheinboldt, On discretization and differentiation of operators with application to Newton's method, *SIAM J. Numer. Anal.* **3** (1966), 143–156.
14. C. R. Rao and S. K. Mitra, "Generalized Inverse of Matrices and Its Applications," Wiley, New York, 1971.
15. N. Shinozaki, M. Sibuya, and K. Tanabe, Numerical algorithms for the Moore–Penrose inverse of a matrix: Direct methods, *Ann. Inst. Statist. Math.* **24** (1972), 193–203.
16. F. Smithies, "Integral Equations," Cambridge Tracts in Math. and Math. Phys., No. 49, Cambridge Univ. Press, London, 1958.
17. O. N. Strand, Theory and methods related to the singular-function expansion ann Landweber's iteration for integral equations of the first kind, *SIAM J. Numer. Anal.* **11** (1974), 798–825.
18. M. Z. Nashed, Ed., "Generalized Inverses and Applications," Academic Press, New York, 1976.